Biophysical Journal Volume 100 June 2011 2736–2744

Measuring the Folding Landscape of a Harmonically Constrained Biopolymer

Michel de Messieres, Barbara Brawn-Cinani, and Arthur La Porta*
Department of Physics, Institute for Physical Science and Technology Biophysics Program, University of Maryland, College Park, Maryland

ABSTRACT Pioneering studies have shown that the probability distribution of opening length for a DNA hairpin, recorded under constant force using an optical trap, can be used to reconstruct the energy landscape of the transition. However, measurements made under constant force are subject to some limitations. Under constant force a system with a sufficiently high energy barrier spends most of its time in the closed or open conformation, with relatively few statistics collected in the transition state region. We describe a measurement scheme in which the system is driven progressively through the transition by an optical trap and an algorithm is used to extract the energy landscape of the transition from the fluctuations recorded during this process. We illustrate this technique in simulations and demonstrate its effectiveness in experiments on a DNA hairpin. We find that the combination of this technique with the use of short DNA handles facilitates a high-resolution measurement of the hairpin's folding landscape with a very short measurement time.

INTRODUCTION

Background

The concept of a transition state, or activation complex, which limits the rate at which chemical reactions proceed, dates to the late 19th and early 20th century (1–3). Later work by Kramers (4) provided an explicit model for chemical reaction kinetics in terms of thermally driven diffusion on an effective energy landscape. The one-dimensional domain in which the diffusion is assumed to occur is the reaction coordinate, a variable with dimension of length that parameterizes the physical rearrangement necessary for the reaction to occur. The height of the landscape is the effective free energy of the system as a function of the reaction coordinate. In this formalism, transition states are peaks in the energy landscape that impede diffusion between the initial and final states along the reaction coordinate.

Kramers was motivated by simple chemical reactions that could be assumed to occur by a well-defined pathway. However, this model has been extremely successful in describing the chemical kinetics of biological macromolecules, which often undergo complex conformational dynamics in the process of folding into their active conformation or catalyzing a chemical reaction. Although the system typically explores a high-dimensional phase space in the course of the transition, it is often possible to characterize it in terms of a well-defined effective path that connects the initial and final conformation (5–8).

The development of single-molecule manipulation techniques (based on optical trapping, magnetic force, atomic force microscopy, and fluorescence) has made it possible to monitor progression of a folding/unfolding transition of a biomolecule, opening a new window into Kramer's model (9,10). One widely used technique employs optical tweezers. The molecule in question is localized and tagged with a microsphere that is held in an optical trap. If the conformational change required to complete the reaction pulls the microsphere away from the center of the trap, the molecule will have to perform an additional portion of work to complete the reaction. Furthermore, the progress of the reaction can be monitored by observing the motion of the microsphere with respect to the trap center. If the optical trap is configured to produce constant force F_0, opposing the motion, and a displacement Δx occurs in the course of the reaction, a term $F_0 \times \Delta x$ is added to the free energy of the system. The effect is to uniformly tilt the energy landscape and modify the transition state and final state energy by a well-defined amount (11). Large forces can also result in an alteration of the precise location of the transition state along the reaction coordinate, depending on the shape of the barrier (12).

There are two methods by which measurements made under an external force have typically been used to characterize kinetics of a folding/unfolding transition. The first is referred to as dynamic force spectroscopy (DFS). In DFS, the height of the barrier and its effective distance from the initial state are determined from the dependence of reaction rate on the applied force, or the dependence of disruption force on the rate at which force is increased (the loading rate). A particularly powerful implementation of this technique is to fit the distribution of disruption forces as a function of loading rate to the predicted functional form (11–15). The results of this analysis, however, are difficult to interpret unless it can be assumed that the reaction rate is limited by a single barrier. Furthermore, the parameters obtained will depend on the assumed shape of the barrier.

Another technique, which is especially useful in studying the folding/unfolding dynamics of biopolymers, is to apply

Submitted September 28, 2010, and accepted for publication March 24, 2011.

*Correspondence: alaporta@umd.edu

Editor: Laura Finzi.

0006-3495/11/06/2736/9 $2.00

doi: 10.1016/j.bpj.2011.03.067

a constant external force $F_{1/2}$ that is sufficient to bring the initial and final conformation of the system to a state of equal occupancy. In this case, the system makes repeated transitions between the two conformations (16) and a histogram of the reaction coordinate—which is defined as the end-to-end extension of the polymer itself—is compiled and used to calculate the probability density function (PDF) defined as $p(x)$ (17,18). The energy as a function of the reaction coordinate x is then obtained using

$$E(x) = -k_B T \ln(p(x)). \quad (1)$$

Knowledge of the shape of the energy landscape connecting the initial and final conformation is sufficient to determine the dependence of the reaction rate on the applied force. An advantage of this technique is that it produces unambiguous results even in the case that the reaction rate is not controlled by a single barrier.

Although the transition rate at zero force is often the quantity most relevant to the biological function of a molecule, it has been argued that the detailed energy landscape is more useful in estimating this rate than a measurement of the transition rate itself as a function of force (19). This is because the transition rate depends not only on the characteristics of the energy landscape, but also on the effective diffusion constant for evolution of the reaction coordinate. The attachment of handles to the structure, normally necessary for single-molecule manipulation studies, can impede movement of the structure in solution and reduce this diffusion constant. As a result, an extrapolation of measured transition rate to zero force can dramatically underestimate the native zero-force transition rate (19). An experimental measurement of the energy landscape combined with an estimate of the diffusion constant obtained through numerical simulations has been suggested as a method of obtaining an accurate estimate of the native zero-force transition rate (19).

Harmonic constraint technique

The question remains as to the most effective method of measuring the energy landscape of a folding transition. Measurement at constant force is illustrated in Fig. 1, *a* and *b*. In Fig. 1, *d* and *e*, the energy landscape for folding of a nucleic acid hairpin structure is represented at zero force, and for the force $F_{1/2}$. The expected PDF of the reaction coordinate is shown in Fig. 1, *g* and *h*. For illustrative purposes, it has been assumed that disruption of the structure requires constant energy per basepair (bp) except for a weak basepair in the stem close to the loop (corresponding to a mismatch) as illustrated in Fig. 1, *a* and *b*, creating a recognizable feature in the energy landscape near the transition state. For sufficiently high transition state energy the hairpin at $F_{1/2}$ is expected to exhibit infrequent transitions between the open and closed conformation. As a result, the time required for the system to reach equilibrium between the open and closed conformation may become impractically long, and the hairpin will spend a very small fraction of its time in the vicinity of the transition state.

However, the energy landscape for a folding transition can also be accurately constructed by taking data with the system subject to force, which depends linearly on the end-to-end

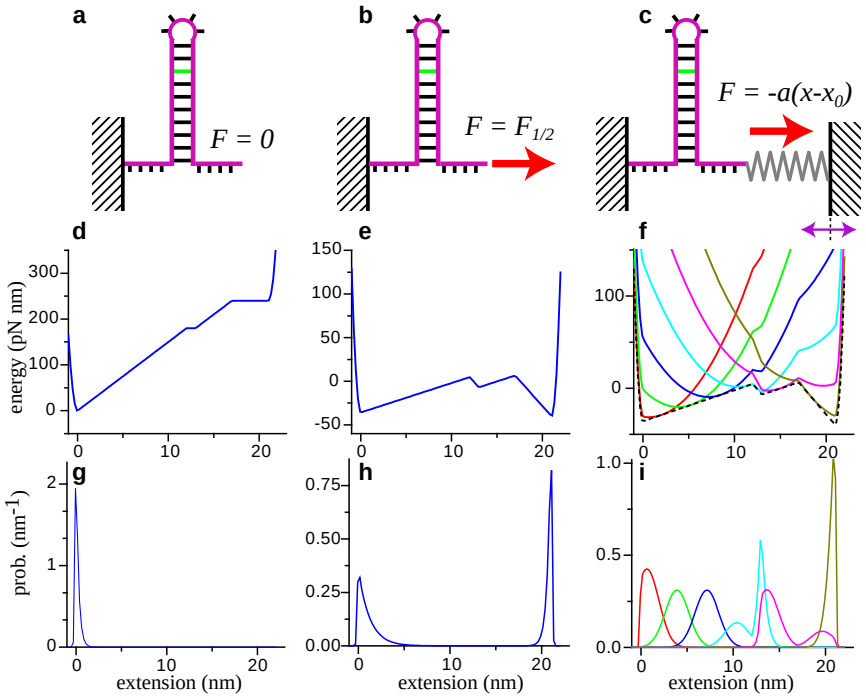

FIGURE 1 (*a–c*) Schematics for experimental configurations in which the hairpin evolves with no external force, for a constant external force $F_{1/2}$ such that the system has equal probability of occupying the open and closed conformations, and subject to a harmonic constraint characterized by stiffness α and origin x_0, which tends to constrain the system along the reaction coordinate x. (*d–f*) Model energy landscapes constructed under the assumption that the energy of the hairpin decreases a constant amount for each basepair hybridized, except for one basepair (*depicted in green*) that hybridizes with no change in energy. Landscapes are shown (*d*) for zero force, (*e*) for force $F_{1/2}$, and (*f*) under harmonic constraint for six different choices of the constraint origin. (*g–i*) Probability density functions of position for the corresponding energy landscapes.

extension of the molecule, so that when the extension increases the tension decreases, and vice versa. Such a harmonic restoring force can be created by taking advantage of the linear-force regime of the optical trap itself, and this experimental configuration has been employed in previous experiments to measure transition rates of hairpins at different constraint positions (20,21). The harmonic constraint manifests itself as a parabolic term superimposed on the native energy landscape of the polymer structure and tends to confine the system to a well-defined region along the reaction coordinate. The native energy landscape in this region can be obtained from the constrained data by correcting for the effect of the imposed constraint. The effect is similar to the umbrella sampling technique, which is frequently used in molecular dynamics simulations to enhance statistical sampling of rarely visited conformations (22).

Although harmonic constraint tends to restrict the system to a limited range of the reaction coordinate, the full landscape can be constructed by taking a series of runs with different origins, confining the system to a set of overlapping regions along the reaction pathway. The effective energy landscapes for six such constraint origins are shown in Fig. 1 *f*. A key feature of this configuration is that if the constraint is sufficiently strong, the large energy barrier that separates the open and closed conformations is no longer the dominant feature of the effective landscape. The predicted PDFs are shown in Fig. 1 *i*, and reflect the parabolic shape of the constraining potential superimposed on the hairpin's native energy landscape. In contrast to the constant-force measurement, every region of the energy landscape is visited with high probability for at least one value of the constraint center, resulting in a much more uniform convergence of statistics. This is in contrast to the PDF obtained from the constant-force landscape, Fig. 1 *h*.

The comparative advantages of measurement at constant force and measurement using a harmonic constraint are apparent in simulations based on the two schemes. We performed Langevin dynamics simulations based on the energy landscapes illustrated in Fig. 1, *e* and *f*, and show the results in Fig. 2. For comparison, the time of the single simulation run at constant force was chosen to be equal to the combined time of the six individual simulations with a harmonic constraint. In Fig. 2 *a* the effective energy landscape at constant force $F_{1/2}$ is contrasted with the landscape for one of the constrained runs restricting the system to remain near the center of the transition. Fig. 2 *b* shows simulated dynamics governed by the two potentials in Fig. 2 *a*. The trajectory at constant force, which is dominated by the large barrier between the open and closed conformations, shows only one transition across the barrier. This time interval has been deliberately selected from the full run of 1200 time units to include the single transition observed. The trajectory that was taken from the constrained system shows rapid fluctuations within the region favored by the constraint potential. The energy barrier that dominates the dynamics at constant force has two distinct disadvantages compared with the constrained landscape. The first is that the position PDF calculated from the simulation at $F_{1/2}$ (Fig. 2 *b*, *black curve*) will show poor convergence in the transition region because the probability of occupying this region is extremely low. The second difficulty is that a system starting in the closed conformation will take a relatively long time to traverse the barrier and sample the open conformation (and vice versa), although thermodynamically, the open and closed conformations should be equally populated. Aside from nonuniform sampling of the reaction coordinate, the system at constant force requires a long time to manifest thermal equilibrium. Even if the system can be sampled at an extremely high rate, it is necessary to wait for the system to equilibrate. The constrained system, in contrast, equilibrates more quickly, facilitating measurement in a short time interval.

Although the individual runs taken with the harmonic constraint converge quickly, no single run samples the full transition. Fig. 2, *c–f*, demonstrates the construction of the

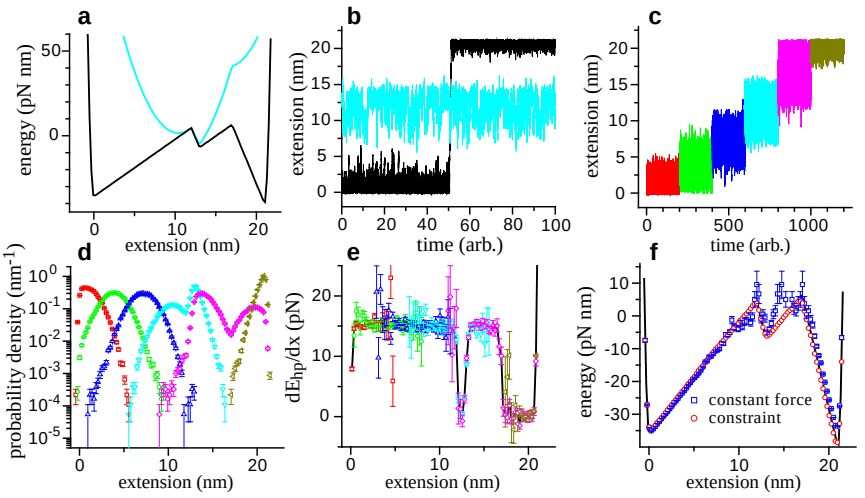

FIGURE 2 (*a*) Comparison of the native energy under constant force (*black*) and for a harmonically constrained force (*cyan*) that constrains the system to remain between the open and closed conformations. (*b*) Simulated trajectory of the hairpin held with constant external force and under the constraining force, corresponding to the two potentials in *a*. The constant-force trace is selected from a run of 1200 time units and the constrained trace is selected from a run of 200 time units. (*c*) Simulated trajectories of the hairpin subject to the harmonic constraint for six positions of the constraint origin. (*d*) Probability density functions calculated from each trajectory in *c* and displayed in matching color. (*e*) dE_{hp}/dx constructed from the data in *d*. The black curve is the combined dE_{hp}/dx curve calculated from Eq. 4. (*f*) Energy landscape calculated from constant-force data (*blue squares*) and harmonic constraint (*red circles*) compared with the effective energy (*black line*), evaluated at $F_{1/2}$.

full energy landscape from a series of runs whose combined domains of convergence span the entire transition. The array of origins used to create the effective landscapes in Fig. 1 *f* is chosen so that the corresponding PDFs show good convergence in overlapping domains of the reaction coordinate, such that every position on the reaction coordinate is well converged in at least one of the PDFs. If we were to apply Eq. 1 to each of these curves we would obtain the total energy, including the contribution of the harmonic constraint. However, if the strength of the constraint α_c and the origin x_c is sufficiently well known, the effect of the constraint can be subtracted to obtain the underlying energy landscape,

$$E_{hp}(x) = -k_B T \ln\left(p(x)\right) - \frac{\alpha_c}{2}(x - x_c)^2 + A_c, \qquad (2)$$

where E_{hp} is the energy of the hairpin we intend to measure and A_c is an arbitrary constant. However, for each value of the constraint position x_c this equation will give $E_{hp}(x)$, which is well converged only within the narrow domain visited by the system. Combining these constrained energy landscapes to form the global energy landscape would require us to choose the constants A_c to form a continuous energy landscape spanning the individual runs. Although this can be done using standard techniques (23), the difficulty of determining the constants A_c is circumvented by calculating the rate of change of E_{hp} with respect to the reaction coordinate,

$$\frac{dE_{hp}}{dx}(x) = -k_B T \frac{d}{dx} \ln\left(p(x)\right) - \alpha_c(x - x_c). \qquad (3)$$

Because the constants A_c have been eliminated, the global dE_{hp}/dx can be obtained by averaging the individual curves, giving each term statistical weight inversely proportional to the square of the standard deviation of the term (24). This is accomplished by evaluating

$$\frac{dE_{hp}}{dx}(x_i) = \frac{\sum_j \left[-k_B T \frac{d}{dx} \ln\left(p_{x_j}(x_i)\right) - \alpha_c\left(x_i - x_j\right) \right] p_{x_j}(x_i)}{\sum_j p_{x_j}(x_i)}, \qquad (4)$$

where x_i is the point at which dE_{hp}/dx is being calculated, x_j is the origin of the harmonic constraint for run j, and $p_{x_j}(x_i)$ is the probability the system is observed at x_i when the constraint origin is x_j. Fig. 2 *d* shows the probability distributions $p_{x_j}(x_i)$, and Fig. 2 *e* shows dE_{hp}/dx calculated for individual constraint origins (symbols with error bars) based on the trajectories in Fig. 2 *c* and the combined dE_{hp}/dx (*black curve*). Note that the individual dE_{hp}/dx curves have large error bars at the margins of their range, but the combined curve is constructed with high resolution throughout, because every part of the reaction coordinate is well converged in at least one of the runs. The function dE_{hp}/dx is a useful way to characterize the transition and can be integrated along x to obtain the energy itself.

Fig. 2 *f* shows the energy landscape (plotted at $F_{1/2}$) obtained from the simulated constant-force measurements and from integration of dE_{hp}/dx for the constrained measurements, compared with the landscape used as the input to the two simulations. The same total amount of measurement time is simulated for both experiments, but the constrained measurement is well converged over the entire landscape, whereas the constant-force data show poor convergence in the critical transition region. Equally good results can, in principle, be obtained from the constant-force configuration, although a ~100-fold longer measurement time is required.

Implications

One might be tempted to conclude from the comparison of the two measurement methods that the constant-force measurement can be made to equal the constrained version as long as the measurement time is extended sufficiently. Although this is true in the idealized conditions of these simulations, in real experiments the constrained method has advantages that cannot generally be duplicated by extending the measurement period of a constant-force data set. One example would be a system that undergoes an irreversible transition. In such a system the constrained measurement could be used to probe the energy landscape in the neighborhood of the irreversible transition, adjusting the constraint to prevent the system from committing to the transition.

More generally, the ability to make measurements in a shorter time interval allows wider latitude in choosing the measurement technology. The long measurement times necessary to reach equilibrium in constant-force measurements have motivated researchers to shun single-beam optical trapping assays in favor of a dual-beam assay (17,25,26). This configuration, in which the construct is stretched between two particles held in two optical traps originating from the same laser, offers unequalled long-range stability, making it possible to record trajectories for many minutes without suffering significant drift. However, the trap centers must be maintained at a distance comparable to the optical wavelength to prevent the two traps from effectively merging into a single trap. The double-stranded DNA handles used to attach the structure are typically chosen to have a total length of ~1700 bp or more to accommodate this geometry (18). Thermally driven fluctuations in the extension of these handles is a source of uncertainty in the measurement of the conformation of the system.

Taking advantage of the fact that the constrained measurements can be completed in a short time interval, we have adopted a single-beam surface-based axial pulling configuration in which the construct is stretched between a surface attachment and the bead. The vertical pulling geometry allows us to use short handles to couple the

nucleic acid construct to the surface and bead, with a total length of 428 bp. This results in a tighter coupling between the structure being studied and the bead, and less uncertainty in the measurement of the system conformation. This is expected to result in higher resolution in the determination of the energy landscape, compared with measurements made with longer handles. Ultimately the choice of handle length will have to take into account other effects, such as the increase in viscous drag that occurs when a bead closely approaches a fixed surface or another bead, or interactions between the bead and the structure being studied. Recently, optical trapping experiments have made use of handles as short as 29 bp offering the possibility that even higher resolution could be achieved (27,28).

MATERIALS AND METHODS

Sample preparation

Two double-stranded DNA handles were generated from PCR using biotin and digoxigenin labeled primers (Invitrogen, Gaithersburg, MD) for attachment to the surface and bead, respectively. Handles were digested with BstXI or BtgI (New England Biolabs, Ipswich, MA) and gel extracted, with final lengths of 234 bp and 194 bp, excluding the 4 bp overhang resulting from the digestion.

Two different hairpins were synthesized (Invitrogen) with a flanking sequence of TTTT on each end for improved ligation efficiency, and 4 base overhangs of cgtg and cgat corresponding to the BtgI and BstXI enzymes, respectively. We will refer to the hairpins as Short (cgtgttttgagtcaa cgtctggatcctgttttcag gatccagacgttgactcttttcgat) and Long (cgtgtttttccgcgcatc tgagtcgaggcagtagccgtcgtctgcga ttttttttcgcagacgacggctactgcctcgactcagatgcg cggttttcgat). (The short hairpin was previously studied in (17)). The handles were ligated to the hairpins, gel extracted, and diluted to ~50 pM in popping buffer (50 mM sodium phosphate buffer pH 7.0, 50 mM NaCl, 10 mM EDTA, and 0.02% Tween 20).

The sample chamber coverslips were scrubbed with Windex, rinsed, and dried. Polyclonal anti-digoxigenin (Roche Molecular Biochemicals, Indianapolis, IN) was diluted to 20 μg/ml in phosphate buffered saline and incubated in the chamber for 20 min. To prevent bead and DNA interactions with the slide, a blotting buffer was incubated in the slide for 3 cycles of 20 min each, containing 1 mg/ml blotting-grade blocker (Bio-Rad, Hercules, CA) in popping buffer. The ligated hairpin constructs were then incubated in the sample chamber for 45 min for attachment of the DNA dig label to the anti-digoxigenin on the surface. Carboxyl microspheres (Bangs Laboratories, Fishers, IN) with radius 410 nm were streptavidin coated, diluted to 0.1 g/ml in popping buffer, and incubated for 20 min for attachment to the biotin labels on the DNA. A final flow through left the sample in popping buffer and oxygen-scavenging solution (721 μg/ml glucose oxidase (Sigma, St. Louis, MO), 144 μg/ml catalase (Sigma), and 3.9 mg/ml glucose) (29). We found that the oxygen-scavenger was essential for working at the high laser powers used.

Instrumentation

Data were collected on a single-beam optical trap using the experimental configuration shown in the Fig. 3 *a* inset. A 1064 nm laser (BL-106C, Spectra-Physics, Santa Clara, CA) was coupled to a single-mode,

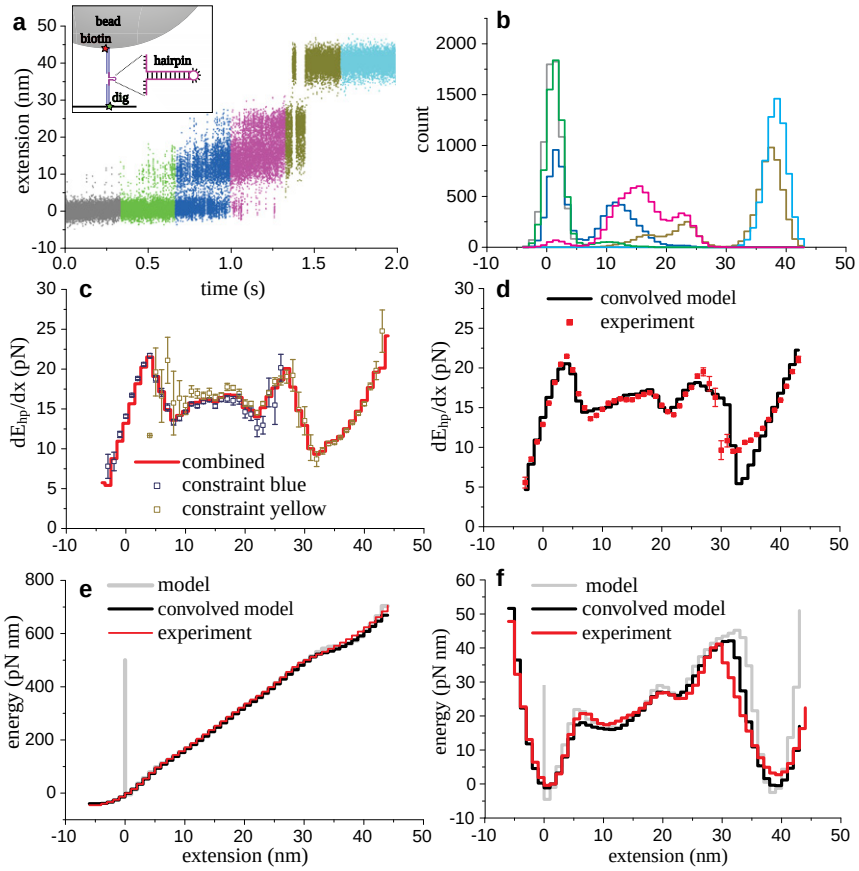

FIGURE 3 Harmonic constraint applied to a single long DNA hairpin. The trap position was moved for 3 scans of duration 2 s. (*a*) Opening distance for the first scan divided into six intervals. Inset shows the experimental configuration, with force applied in the vertical direction. Double-stranded DNA handles with total length 428 bp were used to attach the hairpin between the surface and the trapped bead using digoxigenin and biotin labels, respectively. (*b*) Probability distribution with 1 nm bins for the six intervals of the scan shown in *a*. (*c*) Change in energy/nm shown for two of the intervals (*blue and yellow*) along with the combined result for the six intervals shown in (*red*). (*d*) Change in energy/nm for the combined result of the 3 scans (*red*) compared to the predictions of the convolved model (*black*). (*e*) The model after accounting for single-stranded elasticity (*gray*), the convolved model (*black*), and experimental results (*red*). (*f*) The same data shown in *e*, except plotted at $F_{1/2}$.

polarization-preserving optical fiber (Thorlabs, Newton, NJ). The trap was constructed with a NA 1.49 oil immersion objective (CFI Apo TIRF 100×, Nikon Instruments, Lewisville, TX). Light was collected with a NA 1.4 oil immersion condenser (Nikon Instruments) on a dual-axis position-sensitive diode (DL100-7PCBA3, Pacific Silicon Sensor, Westlake Village, CA). A fraction of the beam incident on the objective lens was also measured to monitor optical power. An acousto-optic modulator (Isomet, Springfield, VA) was used to either stabilize the power of the beam or maintain constant force using software feedback. Data were collected at 60 kHz on a 16-bit digital acquisition board (National Instruments, Austin, TX). The analog signal was filtered using an 8-pole Bessel filter at 30 kHz (Krohn-Hite, Brockton, MA), and filtered offline at 20 kHz. An iris after the condenser clipped the beam to provide a signal dependent on the bead height (30).

Calibration

The double-stranded DNA handles were modeled as a WLC (worm-like chain) (31) with a total contour length of 160 nm determined by assuming 0.338 nm/bp (32) with an additional 15 nm added to estimate the size of the biotin and dig labels, stretch modulus of 1205 pN, and persistence length 43.1 nm (33). Single-stranded DNA was modeled as a WLC using average values from Woodside of persistence length 1.25 nm, contour length per base of 0.625 nm, and neglecting enthalpic contributions (17).

The optical trap was calibrated using standard techniques based on the Fourier spectrum of Brownian fluctuations and scans using a bead rigidly attached to the surface of the sample chamber (34). The position calibrations were adjusted to reproduce the previously reported opening distance of the short hairpin (17). The vertical trap stiffness was measured to be 249 ± 3 pN/μm (standard deviation of the mean) with 860 mW of power on the back of the overfilled objective.

Experimental setup

We studied hairpins in either force-clamp configuration or harmonically constrained configuration. For the force clamp, a software feedback loop controlled the power of the trap through the acousto-optic modulator to maintain a constant force on the hairpin as it fluctuated between the open and closed conformations. For the constrained configuration, we lowered the stage at a constant rate to sweep across the hairpin from the closed conformation to the opened conformation. During constrained scans, the feedback loop controlling the acoustic-optic modulator held the trap power constant. A normalized z signal was calculated as the intensity measured at the detector divided by the intensity incident on the objective (30). We recorded the normalized z signal, trap power, and stage position for the analysis.

Theoretical model

We compared our results to a simple theoretical model for the free energy of the hairpin, based on Mfold v.3.2, (using a monovalent salt correction (35) of 140 mM at 24°C), which gives the energy released by sequential hybridization of the basepairs (36,37). For a given extension of the structure, the number of basepairs hybridized is determined, taking into account the extension of the single-stranded DNA released at the average observed force. The energy is the combination of the hybridization energy specified by Mfold and the elastic energy of the single-stranded DNA released. There is some uncertainty in the increase in extension when the hairpin loop opens. We adjusted this parameter to reproduce the observed well depth for the fully open hairpin. To approximate the blurring effect due to fluctuations of the single-stranded DNA, we applied a pseudoconvolution, in which the width of the Gaussian kernel varies across the landscape by $\sigma = \sqrt{kT/\alpha}$. The varying stiffness α was calculated using the WLC model for the single-stranded DNA released from the hairpin at the average measured force for each bin. This specifies the intrinsic energy landscape of the hairpin.

To account for fluctuations in the apparent extension of the structure arising from thermal fluctuations of the double-stranded DNA handles and trapped bead, we convolved the probability distribution derived from the energy landscape model with a Gaussian of uncertainty 1.1 nm, determined by fitting to the closed well of the hairpin. The model is similar to that employed by Woodside et al., except we account for different parts of the landscape being measured at different average force for constrained measurement. An example of the original model with single-stranded elasticity applied, the convolved model, and experimental data are shown in Fig. 3, *e* and *f*.

RESULTS AND DISCUSSION

Harmonic constraint data analysis

To implement the harmonic constraint method, we lowered the stage at a constant rate averaging 20 nm/s. The resulting records of opening distance versus time were divided into six equally spaced time intervals, each of which appears to be in local equilibrium. Fig. 3 *a* shows the resulting data for one scan of a hairpin. For each interval, the double-stranded DNA handles were approximated as a linear spring using the WLC stiffness at the average force for that interval. The stiffness of the trap was assumed to be constant for all measurements. The trap and the DNA were treated as two linear springs in series with a combined stiffness, $\alpha_c = (1/\alpha_{DNA} + 1/\alpha_{Trap})^{-1}$, where α_c represents the effective stiffness of the harmonic constraint. The energy of the system including the constraint follows Eq. 3, where E_{hp} includes the energy for stretching the single-stranded DNA after it is liberated from the hairpin. A scaling correction (which averaged 2%) was applied to each data set so that the total basepairs released matched the expected value. A set of probability distributions of the measured opening distance was accumulated in equally spaced 1.0 nm bins (Fig. 3 *b*).

A discrete version of Eq. 4 gives us a measurement of dE_{hp}/dx for each interval. As in the simulation, each interval only provides accurate results for one region of the landscape. In Fig. 3 *c* we show two of these intervals (*blue and yellow*) for clarity. Blue provides good statistics in the middle region of the landscape, whereas yellow provides good statistics for the regime when the hairpin is mostly open. The energy landscape for this hairpin is more irregular than that assumed in the simulation, therefore the regions of good statistical convergence are more widely distributed in the experimental data. The combined graph of dE_{hp}/dx was calculated using Eq. 4 and in Fig. 3 *c* red we show the combined result of the six intervals from the first scan. In Fig. 3 *d*, we show the combined results (*red*) for three scans on the same sample in comparison with the predictions of the convolved model (*black*). We integrated the curve in Fig. 3 *d* (*red*) to produce the complete landscape shown in Fig. 3 *e*. For a better comparison of the features in the landscape, we apply an effective force $F_{1/2}$ to tilt the landscapes so that the open and closed wells are occupied with equal probabilities. The value of $F_{1/2}$ for data and the model

deviated by 0.34 pN, which is within the experimental uncertainty of the force calibration.

Force-clamp data analysis

To compare with the harmonically constrained data analysis, we held the same hairpin sample in force-clamp mode and measured data for a time period equal to the total time of the constrained data sweep. The opening distance is plotted over 6 s (Fig. 4 *a*) giving the probability distribution (Fig. 4 *b*). At $F_{1/2}$ (~15 pN) the mean time between transitions is much larger than the proscribed measurement time. To obtain at least one transition during a measurement period, we began with the system in the open conformation and reduced the force to 14.2 pN, which increases the rate of transitions to the closed conformation. The system remained in the open conformation for 14 s before making a transition to the closed conformation and we selected a 6 s interval centered on the transition for analysis. The probability distribution can be compared with the harmonic constraint result (Fig. 3 *b*). We also show the log of the probability with uncertainties (Fig. 4 *c*). The final landscape was reconstructed directly from Eq. 1, correcting for the energy for stretching the single-stranded DNA. (Fig. 4 *d*) Because the system made only one transition, the fraction of time spent in the open and closed conformations is arbitrary and the statistical convergence of the transition state is extremely poor. The shape of the energy landscape in the neighborhood of the open and closed conformations is statistically converged and well defined, but the relative energy of the open and closed conformations and the shape of the transition state landscape are not determined by this data.

Combined samples

In Fig. 5 we show the combined results from five different hairpin samples. Each hairpin contributed one scan of six intervals with a bin spacing of 1 nm. These intervals were combined with appropriate weights by the same method we used to combine intervals for the single hairpin sample in Fig. 3.

Some of our long hairpin samples deviated from the model with an initial opening force significantly below forces required to open the remainder of the hairpin. These features were reproducible on subsequent scans and we attributed them to mutations in the original synthesis of the 99 bp hairpin, contamination or deterioration of the sample, or optical damage. We assumed that any defects would lower rather than increase the hybridization energy and therefore implemented a simple scoring system to delete hairpins that had anomalously low total opening energy. We scored each sample as the highest force measured before it opened beyond 5 bp. The top 40% of samples were found to be self-consistent and we combined them to produce the average landscape shown in Fig. 5.

In Fig. 5, we find good agreement between the measured landscape and the model, which includes the expected blurring resulting from of the double-stranded DNA handles. The loss of resolution compared to the intrinsic landscape of the hairpin is 1.1 nm, which corresponds to approximately 1 basepair opening of the hairpin.

DISCUSSION

The principal benefit of collecting data under a harmonic constraint rather than constant force occurs in cases where

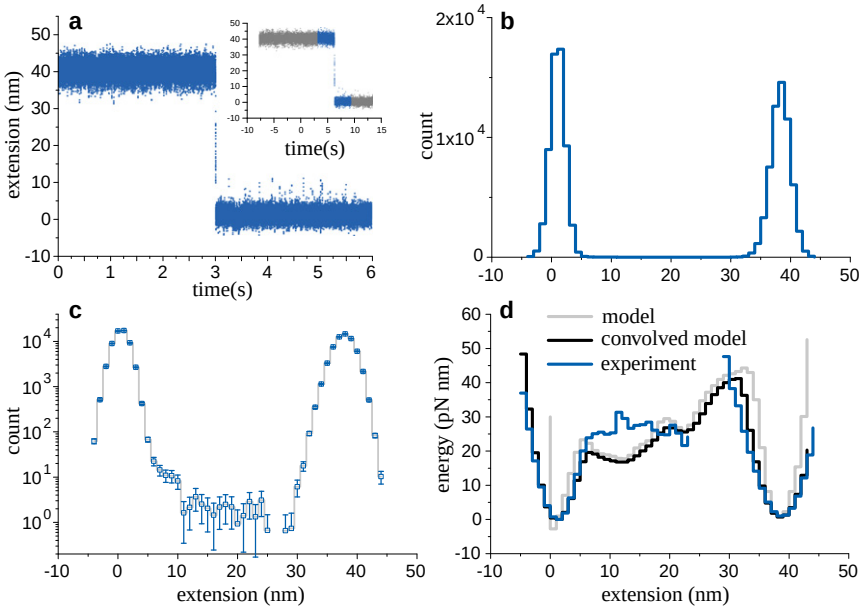

FIGURE 4 Force-clamp data for the same hairpin sample measured in Fig. 3. (*a*) Opening distance measured as the hairpin makes one transition during 6 s (compare to Fig. 3 *a*). Inset: Shows the same data over a longer time frame and the portion selected for analysis. (*b*) Probability distribution of the opening distance with 1 nm bins shown in *a* (compare to Fig. 3 *b*). (*c*) Log of the probability distribution with uncertainties. (*d*) Reconstructed energy landscape (*thick blue line*), the model (*light*), and convolved model (*dark*) (compare to Fig. 3 *f*).

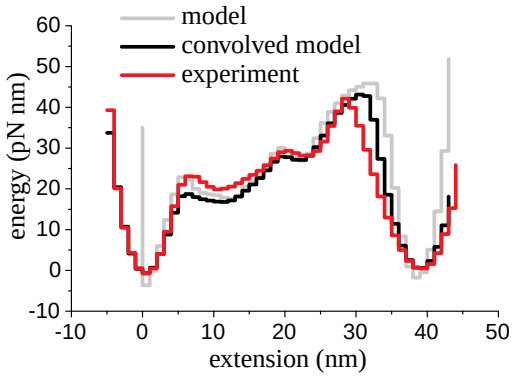

FIGURE 5 Energy landscape (*thick red line*) produced by combining scans of five different data samples, the model (*light*), and convolved model (*dark*). Data analyzed with bin spacing 1.0 nm and six intervals for each scan.

the system has a large energy barrier between the open and closed conformation. Provided the constraint can be made strong enough, the effective barrier will be smaller for the constrained system and it will equilibrate much more quickly under the harmonic constraint than under constant force. Because the occupancy of the transition state is proportional to $\exp(-E_t/k_BT)$, where E_t is the transition state energy, a dramatic decrease in the transition time can be achieved from even a modest reduction in the transition state energy. In such cases the harmonic constraint method allows the landscape to be reconstructed from an ensemble of constrained measurements with good statistical convergence and high resolution using a much smaller total measurement time than would be necessary for a constant-force measurement (compare Fig. 3 *f* and Fig. 4 *d*).

The ability to complete the measurement in a short time reduces the sensitivity to drift, and allows us to use a simplified experimental configuration in which the structure is attached to the trapped bead using a relatively short pair of double-stranded DNA handles. Fluctuations in the length of the handle are a significant source of uncertainty in the instantaneous state of the system and cause blurring of the measured energy landscape. Comparison of our measured landscape with the Mfold-based model suggests that fluctuations of the handle introduce blurring equivalent to convolution of the true landscape with a Gaussian function of standard deviation 1.1 nm. Neglecting end effects, thermal fluctuations in the length of the handles are expected to scale as the square root of the length. Because our handles are a factor of 4 shorter than those used in double-beam experiments by Woodside et. al (17), this implies that our resolution should be better by a factor of 2. Woodside employed deconvolution to compensate for handle effects. However, the landscapes reported by Woodside in the absence of deconvolution show a level of blurring consistent with this estimate. We expect that further reduction in the handle length, either in the vertical pulling optical trap assay or in other single-molecule assays, will further increase the resolution of the measurement of the energy landscape.

CONCLUSION

The method we have presented demonstrates that quantitatively accurate energy landscapes can be obtained from data taken with strong harmonic constraint by removing the effect of the constraint and combining data taken at different constraint positions. We have demonstrated this technique in a single-beam optical trapping experiment, but we anticipate that it can be applied to measurements made using other technologies, including magnetic tweezers or atomic force microscopy, provided that the harmonic constraint can be implemented with sufficient stiffness and the measurement resolution is adequate. The harmonic constraint can facilitate more rapid measurement and may also be preferred in situations where the system is near an effectively irreversible transition. As a result, this technique may broaden the scope of energy landscape reconstruction to systems that currently can only be studied by DFS techniques.

This project was supported by the Maryland Technology Development Corporation.

REFERENCES

1. Arrhenius, S. 1889. On the rate of reaction of the inversion of sucrose by acids. *Z. Phys. Chem.* 4:226–248.
2. Eyring, H. 1935. The activated complex in chemical reactions. *J. Chem. Phys.* 3:107–115.
3. Laidler, K. J., and M. C. King. 1983. The development of transition-state theory. *J. Phys. Chem.* 87:2657–2664.
4. Kramers, H. A. 1940. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica.* VII:284–304.
5. Plotkin, S. S., and J. N. Onuchic. 2002. Understanding protein folding with energy landscape theory. Part I: Basic concepts. *Q. Rev. Biophys.* 35:111–167.
6. Onuchic, J. N., and P. G. Wolynes. 2004. Theory of protein folding. *Curr. Opin. Struct. Biol.* 14:70–75.
7. Wolynes, P. G., J. N. Onuchic, and D. Thirumalai. 1995. Navigating the folding routes. *Science.* 267:1619–1620.
8. Dill, K. A., S. B. Ozkan, …, T. R. Weikl. 2008. The protein folding problem. *Annu. Rev. Biophys.* 37:289–316.
9. Borgia, A., P. M. Williams, and J. Clarke. 2008. Single-molecule studies of protein folding. *Annu. Rev. Biochem.* 77:101–125.
10. Neuman, K. C., and A. Nagy. 2008. Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy. *Nat. Methods.* 5:491–505.
11. Evans, E. 1999. Looking inside molecular bonds at biological interfaces with dynamic force spectroscopy. *Biophys. Chem.* 82:83–97.
12. Dudko, O. K., G. Hummer, and A. Szabo. 2008. Theory, analysis, and interpretation of single-molecule force spectroscopy experiments. *Proc. Natl. Acad. Sci. USA.* 105:15755–15760.
13. Evans, E. 2001. Probing the relation between force–lifetime–and chemistry in single molecular bonds. *Annu. Rev. Biophys. Biomol. Struct.* 30:105–128.

14. Koch, S. J., and M. D. Wang. 2003. Dynamic force spectroscopy of protein-DNA interactions by unzipping DNA. *Phys. Rev. Lett.* 91:028103.

15. Lynch, S., H. Baker, …, K. Sinniah. 2009. Single molecule force spectroscopy on G-quadruplex DNA. *Chemistry.* 15:8113–8116.

16. Liphardt, J., B. Onoa, …, C. Bustamante. 2001. Reversible unfolding of a single RNA molecule by mechanical force. *Science.* 292:733–737.

17. Woodside, M. T., P. C. Anthony, …, S. M. Block. 2006. Direct measurement of the full, sequence-dependent folding landscape of a nucleic acid. *Science.* 314:1001–1004.

18. Woodside, M. T., W. M. Behnke-Parks, …, S. M. Block. 2006. Nanomechanical measurements of the sequence-dependent folding landscapes of single nucleic acid hairpins. *Proc. Natl. Acad. Sci. USA.* 103:6190–6195.

19. Hyeon, C., G. Morrison, and D. Thirumalai. 2008. Force-dependent hopping rates of RNA hairpins can be estimated from accurate measurement of the folding landscapes. *Proc. Natl. Acad. Sci. USA.* 105:9604–9609.

20. Wen, J.-D., M. Manosas, …, I. Tinoco, Jr. 2007. Force unfolding kinetics of RNA using optical tweezers. I. Effects of experimental variables on measured results. *Biophys. J.* 92:2996–3009.

21. Manosas, M., J.-D. Wen, …, F. Ritort. 2007. Force unfolding kinetics of RNA using optical tweezers. II. Modeling experiments. *Biophys. J.* 92:3010–3021.

22. Torrie, G. M., and J. P. Valleau. 1977. Nonphysical sampling distributions in Monte Carlo free-energy estimation: umbrella sampling. *J. Comput. Phys.* 23:187–199.

23. Kumar, S., D. Bouzida, R. H. Swendsen, P. A. Kollman, and J. M. Rosenberg. 1992. The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J. Comput. Chem.* 13:1011–1021.

24. Denesyuk, N. A., and J. D. Weeks. 2009. Equilibrium and nonequilibrium effects in the collapse of a model polypeptide. *Phys. Rev. Lett.* 102:108101.

25. Greenleaf, W. J., M. T. Woodside, …, S. M. Block. 2005. Passive all-optical force clamp for high-resolution laser trapping. *Phys. Rev. Lett.* 95:208102.

26. Shaevitz, J. W., E. A. Abbondanzieri, …, S. M. Block. 2003. Backtracking by single RNA polymerase molecules observed at near-base-pair resolution. *Nature.* 426:684–687.

27. Junier, I., A. Mossa, …, F. Ritort. 2009. Recovery of free energy branches in single molecule experiments. *Phys. Rev. Lett.* 102:070602.

28. Huguet, J. M., C. V. Bizarro, …, F. Ritort. 2010. Single-molecule derivation of salt dependent base-pair free energies in DNA. *Proc. Natl. Acad. Sci. USA.* 107:15431–15436.

29. Landry, M. P., P. M. McCall, …, Y. R. Chemla. 2009. Characterization of photoactivated singlet oxygen damage in single-molecule optical trap experiments. *Biophys. J.* 97:2128–2136.

30. Pralle, A., M. Prummer, …, J. K. Hörber. 1999. Three-dimensional high-resolution particle tracking for optical tweezers by forward scattered light. *Microsc. Res. Tech.* 44:378–386.

31. Marko, J. F., and E. D. Siggia. 1995. Stretching DNA. *Macromolecules.* 28:8759–8770.

32. Saenger, W. 1988. Principles of Nucleic Acid Structure. Springer-Verlag, New York.

33. Wang, M. D., H. Yin, …, S. M. Block. 1997. Stretching DNA with optical tweezers. *Biophys. J.* 72:1335–1346.

34. Svoboda, K., and S. M. Block. 1994. Biological applications of optical forces. *Annu. Rev. Biophys. Biomol. Struct.* 23:247–285.

35. Peyret, N. 2000. Prediction of nucleic acid hybridization: parameters and algorithms. PhD dissertation. Wayne State University, Detroit, MI.

36. Zuker, M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* 31:3406–3415.

37. SantaLucia, Jr., J. J. 1998. A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc. Natl. Acad. Sci. USA.* 95:1460–1465.